

# Entanglement can completely defeat quantum noise

Jianxin Chen,<sup>1,2</sup> Toby S. Cubitt,<sup>3</sup> Aram W. Harrow,<sup>4</sup> and Graeme Smith<sup>5</sup>

<sup>1</sup>*Department of Mathematics & Statistics, U. of Guelph, Guelph, Ontario, Canada*

<sup>2</sup>*Institute for Quantum Computing, U. of Waterloo, Waterloo, Ontario, Canada*

<sup>3</sup>*Departamento de Análisis Matemático, U. Complutense de Madrid, Plaza de Ciencias 3, 28040 Madrid, Spain*

<sup>4</sup>*Department of Computer Science & Engineering, U. of Washington, Seattle, WA 98195, USA*

<sup>5</sup>*IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA*

We describe two quantum channels that individually cannot send any information, even classical, without some chance of decoding error. But together a single use of each channel can send quantum information perfectly reliably. This proves that the zero-error classical capacity exhibits *superactivation*, the extreme form of the superadditivity phenomenon in which entangled inputs allow communication over zero capacity channels. But our result is stronger still, as it even allows zero-error *quantum* communication when the two channels are combined. Thus our result shows a new remarkable way in which entanglement across two systems can be used to resist noise, in this case perfectly. We also show a new form of superactivation by entanglement shared between sender and receiver.

Sending information over a noisy communication channel usually requires error correction. The best transmission rate possible, optimized over all conceivable error-correction strategies, is called the capacity of the channel. The capacity tells us the value of a noisy channel for communication and is measured in bits per channel use. Capacities are central to the theory of information initiated by Shannon [1], and serve as guideposts for the development of practical communication schemes.

The usual setting for information theory is the asymptotic regime, where sender and receiver have many independent uses of a fixed noisy channel. The probability of transmission error is required to vanish in the limit of many channel uses, and the resulting capacity is called the Shannon capacity. A more demanding requirement is to insist that the error probability be *exactly* zero. This leads to zero-error information theory, also studied by Shannon [2]. The zero-error setting has a more combinatorial flavor; indeed, a large part of modern graph theory owes its origins to the study of zero-error communication [3]. Zero-error information theory is most relevant when the asymptotic guarantees of Shannon theory are insufficient—either because the number of channel uses isn't large enough to make the probability of error small, or because absolutely no error can be tolerated. Furthermore, it is related to the rate at which the error probability tends to zero in the usual Shannon capacity [4].

Ultimately, noisy communication links are described by quantum mechanics, and in systems such as optical communication, quantum effects cannot be neglected. When considering the zero-error capacity of quantum channels, we may consider either the classical or quantum capacities, measuring respectively the rate at which a channel may send bits or qubits without any error. The resulting coding problems lead to rich generalizations of the graph theory problems arising from classical channels [5].

Even classically, the zero-error capacity is quite different from the Shannon capacity. For example, it is

non-additive [2]. However, some basic properties are common to both capacities. One of the most basic is the behavior of zero-capacity channels, i.e. channels that are too noisy to transmit any information. It seems like combining two such completely useless channels will still not allow communication. Indeed, for classical channels, this intuition is correct. The only classical channels with no zero-error capacity are those where every pair of inputs have some non-zero probability of being confused at the output (otherwise we could use a non-confusable pair of inputs to send one bit). But if we use two such channels together, any pair of inputs to the combined channel can also be confused, so the joint channel has zero capacity.

Remarkably, we show that this elementary property of classical channels fails for quantum channels: there are pairs of quantum channels, each with no classical zero-error capacity at all, yet which *do* allow perfect classical transmission when the two channels are combined. This striking phenomenon is known as *superactivation* [6], and to our knowledge this is the first superactivation of a classical capacity of standard quantum channels. In fact, we can strengthen this result to show that the joint channel can even transmit far more delicate *quantum* information with zero error, so these channels superactivate both the classical and quantum zero-error capacity *simultaneously*, an extreme form of superactivation never before seen.

It is a little like having two pipes, both completely blocked, allowing nothing to flow through them. Yet, by plumbing the blocked pipes in parallel, water can flow. Of course, this analogy breaks down due to quantum effects. Indeed, entanglement is at the heart of this remarkable superactivation phenomenon. Our results show that using input states entangled across the two inputs to the joint channel, we can completely defeat noise, allowing perfect transmission where none would otherwise be possible. Finally, we also show that entanglement can be used to completely defeat noise in a new form of superactivation: superactivation by entanglement, where entanglement

between the sender and receiver allows perfect communication with a zero-capacity channel.

*Related phenomena.* Superactivation has previously been found only for the quantum capacity of quantum channels [6]. By contrast, the classical capacity is nonzero for any nontrivial quantum channel, so superactivation is trivially impossible in the standard Shannon setting. There, the possibility of superadditivity (two channels having greater asymptotic capacity when combined) remains a major open question, since the additivity violation of Hasting [7] only addressed one-shot (Holevo) capacities. Private communication is intermediate between classical and quantum communication. Here, superadditivity has been observed [8], while superactivation remains an open question. In the zero-error case, other researchers have found activation results for single copies of quantum channels [9]. It has also been shown that the zero-error capacity even of classical channels can be increased by shared entanglement between sender and receiver, though it cannot be superactivated [10]. Recently, others have started to explore how quantum zero-error capacities relate to the graph-theoretic quantities that provide bounds on zero-error capacities of classical channels [5].

*Overview of technical contributions.* We use two key technical ideas. The first is to choose our channels as randomly as possible, subject to certain constraints. The constraints guarantee that combining the two channels allows information to be transmitted. Meanwhile, the random choice helps ensure that the individual channels are noisy, so transmit very little information. This *constrained-randomness* strategy has had many applications in quantum information. Examples include [11], which considered random states subject to a rank constraint; [7, 12], which chose random channels subject to a constraint ensuring their product gave one large output eigenvalue; or the variant in [13], which choose random channels subject to a constraint that guaranteeing a singular output for an appropriate entangled input.

This strategy is very successful in showing that there is a nonzero probability of picking a channel that is noisy enough on a single copy. This is sufficient for all the above results, as they only concern “one-shot” quantities: properties of a single copy of a channel. But we need something stronger; we require *arbitrarily many* copies of the *same* channel to be so noisy that the channel has no zero-error capacity (even asymptotically). We only get to exploit a finite amount of randomness (in the choice of one copy of the channel), yet we want this finite amount of randomness to control a property of an arbitrarily large and highly correlated object (the capacity of arbitrarily many copies of that same channel). However large the probability of picking a suitable channel for a single copy, unless that probability is 1 it will shrink to zero on a growing number of copies. So, on its own, this strategy completely fails to give results for asymptotic quantities such as channel capacities.

Our second technical tool is a new method of controlling the behavior of an unbounded number of copies of the channel through randomness on a single copy using algebraic geometry. Such arguments show that certain bad sets (say, the set of channels for which  $k$  copies can send a classical bit with zero error) have zero measure, so that even a union of countably many of them does as well. In other words, we show that the probability of picking a suitable constrained random channel is exactly 1, thus avoiding any decay in the probability for growing numbers of copies. These techniques rely on a greater knowledge of the structure of the problem, but are still highly general, and we suspect they will have further application in quantum information, including problems in which small errors are tolerated.

*Proof of main result.* We now describe the proof of zero-error superactivation. Recall that two quantum states  $\rho, \sigma$  are perfectly distinguishable exactly when they are orthogonal ( $\text{Tr}[\rho\sigma] = 0$ ). Thus, the classical zero-error capacity of a channel  $\mathcal{E}$  is 0 exactly when no pair of inputs gives orthogonal outputs. Mathematically, we require

$$\forall \psi, \varphi \quad \text{Tr}[\mathcal{E}(\varphi)\mathcal{E}(\psi)] \neq 0. \quad (1)$$

Let  $\circ$  denote composition, define  $\mathcal{E}^*$  by  $\text{Tr } A\mathcal{E}(B) = \text{Tr } \mathcal{E}^*(A)B$  and  $\mathcal{N} := \mathcal{E}^* \circ \mathcal{E}$ . Then Eq. (1) is equivalent to requiring  $\forall \psi, \varphi \quad \text{Tr}[\varphi\mathcal{N}(\psi)] \neq 0$ . This in turn is equivalent to insisting that the (CP, but not necessarily trace preserving) map  $\mathcal{N}$  always has full rank output. This condition was previously used in [13] to find multiplicativity violations for the minimum output rank. Following [13], we can rewrite this condition in terms of the Choi matrix [14]  $\sigma_{AB}$  of the composite map  $\mathcal{N} = \mathcal{E}^* \circ \mathcal{E}$  (recalling that the action of the map can be recovered from the Choi matrix via  $\mathcal{N}(\rho) = \text{Tr}[\sigma_{AB} \cdot (\rho^T \otimes \mathbb{1})]$ ), to obtain  $\text{Tr}[\sigma_{AB} \cdot (\psi \otimes \varphi)] \neq 0$ . In other words, the support of  $\sigma_{AB}$  (denoted  $S_{AB}$ ) contains no product states. [15] The same argument holds for any number of copies  $k$  of the channel. So, for a channel to have no zero-error capacity even asymptotically,  $S_{AB}^{\otimes k}$  must contain no product states for *any* tensor power  $k$  (unlike [13], where  $k = 1$  sufficed).

Furthermore, in contrast to [13], it is no longer true that *any* bipartite subspace  $S_{AB}$  will suffice; the fact that the subspace must now support the Choi matrix of a composite map  $\mathcal{N} = \mathcal{E}^* \circ \mathcal{E}$  imposes extra symmetry requirements on  $S_{AB}$ . It is easy to verify that  $S_{AB}$  must have the following additional properties: (i)  $\mathbb{F}(S_{AB}) = S_{AB}$ , where  $\mathbb{F}(\sum_{ij} \alpha_{ij} |i\rangle |j\rangle) = \alpha_{ij}^* |j\rangle |i\rangle$  swaps the two systems and takes complex conjugates, (ii)  $S_{AB}$  contains a state  $|\psi\rangle_{AB} = \sum_{i,j} \alpha_{ij} |i\rangle_A |j\rangle_B$  whose matrix of coefficients  $M = [\alpha_{ij}]$  is positive-definite (has strictly positive eigenvalues). These two conditions are also sufficient, in the following sense: given a subspace  $S_{AB}$  satisfying (i) and (ii), one can always construct a Choi matrix  $\sigma_{AB}$  supported on  $S_{AB}$  which corresponds to some map  $\mathcal{N} = \mathcal{E}^* \circ \mathcal{E}$ . To see this, choose a (not necessarily orthonormal) basis

$|\psi_k\rangle$  for  $S_{AB}$  whose coefficient matrices  $M_k$  are positive-definite (condition (ii) guarantees that such a basis exists). Denoting the eigenvectors of  $M_k$  by  $|\phi_i^k\rangle$ , the matrix

$$\rho_{AB} = \sum_{ijk} |\phi_i^k\rangle_A |k, i\rangle_B \langle \phi_j^k|_A \langle k, j|_B. \quad (2)$$

is (up to rescaling) a Choi matrix for a channel  $\mathcal{E}$ , such that the Choi matrix of  $\mathcal{N} = \mathcal{E}^* \circ \mathcal{E}$  is supported on  $S_{AB}$ .

To prove superactivation, we need a *pair* of channels  $\mathcal{E}_1$  and  $\mathcal{E}_2$ , each satisfying Eq. (1) so that it has no zero-error capacity, but such that the joint channel  $\mathcal{E}_1 \otimes \mathcal{E}_2$  *does* have positive capacity. For the latter, we need a pair of input states that are mapped to orthogonal outputs by the joint channel, i.e.

$$\exists \psi, \varphi \quad \text{Tr}[(\mathcal{E}_1 \otimes \mathcal{E}_2)(\psi) \cdot (\mathcal{E}_1 \otimes \mathcal{E}_2)(\varphi)] = 0, \quad (3)$$

as we can use these states to perfectly transmit 1 bit. Generalising [13], we choose these inputs  $\psi, \varphi$  to be maximally entangled states  $|\omega\rangle = \sum_i |i\rangle |i\rangle / \sqrt{d}$  and  $|\omega'\rangle = (X \otimes \mathbb{1})|\omega\rangle$ , where  $X$  is the unitary consisting of 1's down the anti-main diagonal (i.e. the generalisation to arbitrary dimension of the Pauli matrix  $\sigma_x$ ). Rewriting Eq. (3) in terms of the Choi matrices  $\sigma_{1,2}$  of the composite maps  $\mathcal{N}_{1,2} = \mathcal{E}_{1,2}^* \circ \mathcal{E}_{1,2}$ , we obtain for this choice of input states that the Choi matrices must satisfy  $\text{Tr}[\sigma_1^T \cdot (X \otimes \mathbb{1}) \sigma_2 (X \otimes \mathbb{1})] = 0$ . But, denoting the supports of  $\sigma_{1,2}$  by  $S_{1,2}$ , this simply states that the subspaces  $S_1$  and  $(X \otimes \mathbb{1})S_2$  should be orthogonal. We might as well take  $S_2 = (X \otimes \mathbb{1})S_1^\perp$ , since this still allows zero-error communication with the composite channel, while making  $S_2$  as large as possible can only help suppress the single-use capacity.

Our task now reduces to finding an appropriate  $S_1$ , which we call simply  $S$  from now on. To summarize the constraints described above, we require:

- (i)  $(S^{\otimes k})^\perp$  contains no product states for any  $k$ ;
- (ii)  $((S^\perp)^{\otimes k})^\perp$  contains no product states for any  $k$ ;
- (iii)  $\mathbb{F}(S) = S$ ;
- (iv)  $\mathbb{F}((X \otimes \mathbb{1}) \cdot S) = (X \otimes \mathbb{1}) \cdot S$ ;
- (v)  $S$  contains a state with positive-definition coefficient matrix;
- (vi)  $(X \otimes \mathbb{1}) \cdot S$  contains a state with positive-definition coefficient matrix.

Properties (iii)–(vi) guarantee that  $S_1 = S$  and  $S_2 = (X \otimes \mathbb{1})S^\perp$  correspond to valid channels. Our choice of  $S_2$  ensures that these channels together can communicate one bit without error. Most of the remaining work is showing that a random  $S$  satisfies (i)–(ii): arbitrary tensor powers contain no product states, ensuring that the individual channels have no zero-error capacity. (In stating property

(ii), we have used the fact that the set of product states is left invariant by  $(X \otimes \mathbb{1})$ .) A priori, this appears extremely demanding, since we must satisfy an infinite number of constraints simultaneously; indeed, we only get to choose a subspace from a constant number of dimensions, but we need to rule out product states on an unbounded number of tensor copies. However, algebraic geometry arguments will show, remarkably, almost *all* subspaces (all but a measure-zero set) satisfy properties (i)–(ii).

There is a standard way to represent a subspace as a vector [16], writing  $S$  as the antisymmetric product of an orthonormal basis of  $S$  (e.g. consider the unique state of  $\dim S$  fermions with state space  $S$ ). With this parameterisation, we can see that the set  $E_k$  of all subspaces whose  $k^{\text{th}}$  tensor power *does* contain a product state is given by a set of simultaneous homogenous polynomial equations [17]. These are the subspaces that we *don't* want; we are looking for a subspace that is *not* in any  $E_k$ . In general, the set of zeros of a set of polynomials will either have measure 0, or will comprise the entire space (Fig. 1) [18].

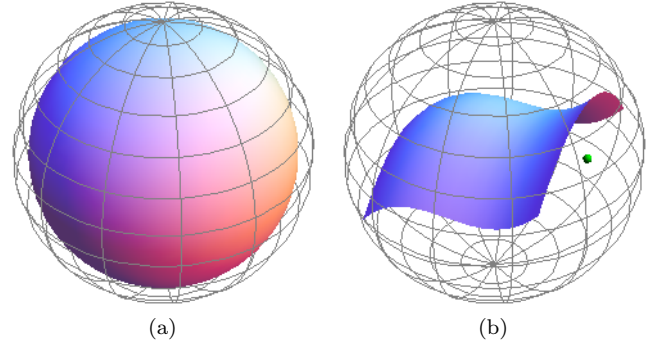


FIG. 1: The set of zeros of a set of simultaneous polynomial equations either (a) comprises the entire space, or (b) has measure 0. To show that it is measure 0, it therefore suffices to find a single point outside of the set (b).

Therefore, to show that  $E_k$  has zero measure, it suffices to show that there is a *single* point outside of it, thereby ruling out the possibility that it is the entire space. To do this we construct such a subspace from an unextendible product basis (UPB): a basis on a bipartite space which cannot be extended by adding any further orthogonal product states. Now, the orthogonal complement of the span of a UPB is by definition a subspace that contains no product states and, since the tensor product of two UPBs is again a UPB [17], this is also true for any tensor power of this subspace, as required. This subspace will certainly not satisfy the other requirements (ii)–(vi) (in particular, its orthogonal complement is just the span of the UPB, which clearly *does* contain product states, dramatically failing to satisfy (ii)). But its existence shows that there

is a subspace that is not contained in any  $E_k$ , which is sufficient to collapse each  $E_k$  to a set of zero measure (Fig. 1). Thus  $\cup_{k \geq 1} E_k$  has zero measure, so property (i) holds for a random  $S$  with probability 1. Since  $S^\perp$  is also uniformly random, property (ii) automatically holds with probability 1 as well.

Our argument must be refined to handle (iii–vi). Start with (iii) and (iv), which are linear constraints on the subspace, and thus translate into polynomial constraints in the coordinates parameterising the subspace. Since the intersection of the solutions to two sets of simultaneous polynomial equations is the set defined by the union of both sets of polynomials, we can use the preceding argument *within* this intersection. The only modification is that we now choose a UPB that satisfies the symmetry requirements. This is achieved by symmetrizing an arbitrary UPB. Since adding additional product states to a UPB maintains the UPB property, this requires only that the initial UPB is not too large, which holds already for the UPBs from [19]. Therefore, we can choose a random subspace satisfying requirements (iii) and (iv) and with probability 1 it will also satisfy requirements (i) and (ii).

Finally, we address (v) and (vi). These are not algebraic, so we cannot repeat the algebraic geometry argument. However, the requirement that our bipartite subspace contain a state with positive-definite coefficient matrix is quite mild. In particular, if a subspace has this property and we perturb it by a sufficiently small amount, then the positive-definite element stays positive definite. So, around every such subspace there is an open ball of subspaces that also satisfy the positivity requirement. Therefore, the set of subspaces satisfying (v) and (vi) has positive measure, even relative to the set of subspaces satisfying (iii) and (iv). We have seen that the set of subspaces satisfying (i) and (ii) is full measure within the set of subspaces satisfying (iii) and (iv). The intersection of a positive-measure set with a full-measure set has positive measure, so the set of subspaces satisfying (i) to (vi) has positive measure. So, at least one such subspace  $S$  must exist. We have already shown that this is equivalent to the existence of a channel that superactivates the classical zero-error capacity, so we are done.

Armed with these techniques, we can extend our result to show the joint channel can even transmit *quantum* information with zero error [20]. Thus, we want the joint channel to transmit at least one qubit perfectly, meaning that some two-dimensional subspace is transmitted undisturbed. For this, it is sufficient [20] to find *two* different pairs of orthogonal input states in the same two-dimensional subspace are mapped to orthogonal output states. This just adds another algebraic symmetry condition on  $S$ , so we can deal with it exactly as before.

*Application to entanglement-assisted capacity.* We have shown that by encoding the information into states  $|\omega\rangle$  and  $(X \otimes \mathbb{1})|\omega\rangle$  which are entangled across the two inputs to the joint channel, we can completely defeat

the noise in the channel, even though we couldn't send *any* information through either channel on its own. Entanglement can be used to completely defeat noise in another way, by sharing the entanglement between sender and receiver. To see this, note that we have a channel  $\mathcal{E}_1$  above with no zero-error capacity, even asymptotically, but for which inputs  $|\psi\rangle = |\omega\rangle$ ,  $|\varphi\rangle = (X \otimes \mathbb{1})|\omega\rangle$  to the joint channel  $\mathcal{E}_1 \otimes \mathcal{E}_2$  give orthogonal outputs:  $\text{Tr}[\mathcal{E}_1 \otimes \mathcal{E}_2(\psi) \cdot \mathcal{E}_1 \otimes \mathcal{E}_2(\varphi)] = 0$  (Eq. (3)). But applying a channel cannot increase the orthogonality of two states, so  $\mathcal{E}_1 \otimes I(\psi)$  and  $\mathcal{E}_1 \otimes I(\varphi)$  are orthogonal. Therefore, if the sender and receiver share the maximally entangled state  $|\omega\rangle$ , then the sender can communicate one bit perfectly to the receiver, by either sending her half of the entangled state directly through the channel, or first applying the local unitary  $X$  to her half of the state before sending it. Since the resulting states  $\mathcal{E}_1 \otimes I(\psi)$  and  $\mathcal{E}_1 \otimes I(\varphi)$  are orthogonal, they can be perfectly distinguished by the receiver, thereby transmitting one bit with zero error.

*Acknowledgments.* JC is supported by NSERC. TSC was supported by a Leverhulme early career fellowship, the EU (QAP, QESSENCE, MINOS, COMPAS, CO-QUIT, QUEVADIS), and Spanish grants QITEMAD, I-MATH, and MTM2008-01366. AWH was funded by NSF grants 0916400, 0829937, 0803478, DARPA QuEST contract FA9550-09-1-0044 and the IARPA MUSIQ and QCS contracts. GS is supported by the DARPA QuEST contract HR0011-09-C-0047.

- 
- [1] C. E. Shannon, Bell System Tech. Jnl. **27**, 379 (1948).
  - [2] C. E. Shannon, IRE Trans. Inform. Theory **IT-2**, 8 (1956).
  - [3] J. Körner and A. Orlitsky, IEEE Trans. Inform. Theory **44**, 2207 (1998).
  - [4] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, Inform. Contr. **10**, 65 (1967).
  - [5] R. Duan, S. Severini, and A. Winter, (2010), arXiv:1002.2514; S. Beigi, Phys. Rev. A **82**, 010303(R) (2010).
  - [6] G. Smith and J. Yard, Science **321**, 1812 (2008), arXiv:0807.4935; G. Smith, J. A. Smolin, and J. Yard, (2011), arXiv:1102.4580.
  - [7] M. B. Hastings, Nature Physics **5** (2009), arXiv:0809.3972.
  - [8] K. Li, A. Winter, X. Zou, and G. Guo, arXiv:0903.4308 (2009); G. Smith and J. Smolin, arXiv:0904.4050 (2009).
  - [9] R. Duan, arXiv:0906.2527 (2009).
  - [10] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, arXiv:1003.3195 (2010), (to appear in IEEE Trans. Inform. Theory); Phys. Rev. Lett. **104**, 230503 (2010).
  - [11] P. Hayden, D. W. Leung, and A. Winter, Comm. Math. Phys. **265**, 95 (2006), arXiv:quant-ph/0407049.
  - [12] A. J. Winter and P. Hayden, Comm. Math. Phys. **284**, 263 (2008), arXiv:0807.4753.
  - [13] T. Cubitt, A. W. Harrow, D. Leung, A. Montanaro, and A. Winter, Comm. Math. Phys. **284**, 281 (2008), arXiv:0712.3628.

- [14] The Choi matrix of a map  $\mathcal{N}$  is  $\sigma_{AB} = (\mathcal{N} \otimes \mathbb{1})(|\omega\rangle\langle\omega|)$ , where  $|\omega\rangle = \sum_i |i\rangle|i\rangle/\sqrt{d}$  and  $\mathbb{1}$  is the identity map.
- [15] Ref [13] refers to channels, but the arguments extend to  $\mathcal{N} = \mathcal{E}^* \otimes \mathcal{E}$ , even if it is not a channel. But the arguments go through unchanged for arbitrary linear maps.
- [16] This is known as the *Plücker embedding*.
- [17] T. S. Cubitt, J. Chen, and A. W. Harrow, (2009), arXiv:0906.2547.
- [18] This is because if the zero set of a polynomial contains a set with non-empty interior, then by Taylor expansion that polynomial must be zero everywhere.
- [19] R. Bhat, Int. J. Quant. Inf. **4**, 325 (2006).
- [20] T. S. Cubitt and G. Smith, arXiv:0912.2737 (2009).